10 | 经验：都有哪些高效的SRE组织协作机制？

2020-04-08 赵成

SRE实战手册 进入课程 >

讲述：赵成
时长 13:06　大小 12.00M

你好，我是赵成，欢迎回来。

上节课，我们讲了在互联网企业典型的 SRE 团队中，一般包含几个关键角色，PE、工具开发和稳定性开发，他们分别承担的职责就是业务运维、建设运维自动化平台和建设稳定性平台。在建设这些平台的过程中，这些角色对内要与中间件团队合作，基于微服务和分布式的架构来开发各类平台；对外，还要与业务开发配合，将提升效率和稳定性的能力提供出去。

但是仅仅有组织架构，有了队形还是不够的，各个团队和角色之间必须要配合协作起来，才能发挥出 SRE 的作用，特别是对外与业务开发的合作，这样才能是一个有机的整体。

那怎样才能将这些角色有效地组合到一起呢？今天我就和你分享下我的经验，总结起来就是四个字：**以赛带练**。

什么是“以赛带练”？

“以赛代练”这个术语最早也是在一些体育赛事中提出来的，完整解释是通过比赛**带动**训练。比如足球、篮球或田径等比赛，目的就是让运动员在真实的高强度竞争态势和压力下，充分暴露出个人和团队的不足及薄弱点，然后在后续的训练中有针对性地改进，这样对于比赛成绩的提升有很大帮助，而不是循规蹈矩地做机械性训练。你也注意到，我用的是“带”，而不是“代”，所以整个过程不是用比赛代替训练，它更主要的作用是带动。

同样，这样的策略放在我们的系统稳定性建设中也非常适用。你也可以选择类似的真实且具备高压效果的场景，来充分暴露我们的稳定性存在哪些问题和薄弱点，然后有针对性地进行改进。

类似的场景，比如说极端的海量用户访问，像电商类产品的双十一大促、社交类产品在春晚的抢红包，以及媒体类产品的突发热点新闻等等，用户瞬时访问的流量是极大的，对于系统的稳定性冲击和挑战也就非常大。再或者，是极端的故障场景，比如机房断电、存储故障或核心部件失效等等。

所以，我们讲稳定性建设，一定也是针对极端场景下的稳定性建设。

不知道你有没有发现，其实我们有很多的稳定性保障技术和理念，比如容量规划和压测、全链路跟踪、服务治理、同城双活甚至是异地多活等，基本都是在这些场景下催生出来的，甚至是被倒逼出来的。

如果针对这样的极端场景，我们都可以从容应对，那你可以想象一下，日常一般的问题或小故障处理起来自然也就不在话下了。而且，目前各大电商网站的大促已经基本日常化，每月都会有，甚至每周每天都会有不同的促销活动，所以在这种密集程度非常高的大促节奏下，“以赛带练”的效果就会更加突出。

可以看到，“以赛带练”的目的就是要检验我们的系统稳定性状况到底如何，我们的系统稳定性还有哪些薄弱点。

那么，如果你在自己的团队来落地“以赛带练”的话，**一定要先考虑自己的业务系统应对的极端场景到底是什么，然后基于这些场景去设计和规划。**

SRE 的各个角色如何协作?

赛场有了，要打好这场仗，我们得保证上赛场的选手们能够通力协作。对于我们想要达成的稳定性目标来说，就是要有一套高效的 SRE 协作机制。接下来我还是以电商大促这个场景为例，跟你分享我们 SRE 团队中各个角色的分工，以及可以通过哪些组织形式把各个角色组织起来。

想要把电商大促这场赛事做漂亮，保障准备工作至关重要。这个准备工作前后共有四个关键步骤，我们就一步一步来看。

第一步，大促项目开工会。

在这个会议上，我们会明确业务指标，指定大促项目的技术保障负责人，通常由经验丰富的业务技术成员或平台技术成员承担，同时会明确技术团队的分工以及各个团队的接口人，然后根据大促日期，倒排全链路压测计划。分工和计划敲定了，接下来就是一步步执行了。

第二步，业务指标分解及用户模型分析评审会。

业务指标分解和用户模型分析阶段，需要业务开发和 PE 团队共同配合，主要是共同分析出核心链路，同时 PE 要分析链路上的应用日常运行情况，特别是 QPS 水位，这里就要利用到我们前面 03 讲中示例的 SLO 的 Dashboard，结合这两点，大致判断出要扩容的资源需求。

这里你可能会有疑问，为什么业务开发和 PE 要一同分析？因为每个 PE 是要负责一整条核心链路上的应用的，所以 PE 可以识别出更全局的链路以及整条链路上的容量水位情况。而每位业务开发往往只专注在某几个应用上，所以他可以把业务目标拆解得更细致，并转化成 QPS 和 TPS 容量要求，然后分配到一个个具体应用上，并且每个应用的 Owner 在后面要确保这些应用能够满足各自的 QPS 和 TPS 容量要求。

从这个过程中，你会发现，**PE 会更多地从全局角度关注线上真实的运行状态**，而业务开发则根据这些信息做更细致的分析，甚至是深入到代码层面的分析。

同时，业务开发和 PE 在分析的时候，就要使用到稳定性开发团队提供的全链路跟踪和监控平台了，而不是靠手工提取日志来做统计分析。

接下来，根据容量评估的结果，由 PE 准备扩容资源，然后业务开发的应用 Owner，通过运维自动化平台，做完全自动化的应用部署和服务上线即可，整个过程无需 PE 介入。

第三步，应急预案评审会。

在预案准备阶段，仍然是 PE 与业务开发配合，针对核心链路和核心应用做有针对性的预案讨论。这时就要细化到接口和方法上，看是否准备好限流、降级和熔断策略，策略有了还要讨论具体的限流值是多少、降级和熔断的具体条件是怎样的，最后这些配置值和配置策略都要落到对应的稳定性配置中心管理起来。

这里 **PE 更多地是负责平台级的策略**，比如，如果出现流量激增，首先要做的就是在接入层 Nginx 做限流，并发 QPS 值要设定为多少合适；再比如缓存失效是否要降级到数据库，如果降级到数据库，必然要降低访问 QPS，降低多少合适，等等。

而**业务开发则要更多地考虑具体业务逻辑上的策略**，比如商品评论应用故障，是否可以直接降级不显示；首页或详情页这样的核心页面，是否在故障时可以完全实现静态化等等。

第四步，容量压测及复盘会。

在容量压测这个阶段，就需要 PE、业务开发和稳定性平台的同学来配合了。业务开发在容量规划平台上构造压测数据和压测模型，稳定性平台的同学在过程中要给予配合支持，如果有临时性的平台或工具需求，还要尽快开发实现。

压测过程会分为单机、单链路和全链路几个环节，PE 和业务开发，在过程中结合峰值数据做相应的扩容。如果是性能有问题，业务开发还要做代码和架构上的优化，同时，双方还要验证前面提到的服务治理手段是否生效。

压测过程中和每次压测结束后，都要不断地总结和复盘，然后再压测验证、扩容和调优，直至容量和预案全部验证通过。这个过程一般要持续 2～3 轮，时间周期上要 3～4 周左右。整个过程就会要求三个角色必须要非常紧密地配合才可以。

到这里，整个保障准备过程就结束了。你可以先思考下，这个过程中每个角色发挥的作用有哪些。接下来，我跟你一起提炼总结下。

第一，PE 更加关注线上整体的运行状态，所以视角会更全局一些，业务开发则更关注自己负责的具体应用，以及深入到代码层面的分析工作。

第二，PE 会主要负责平台级的公共部件，如缓存、消息和文件等；DBA 负责数据库，所起到的作用与 PE 相同，他们关注这些平台部件的容量评估、服务治理策略，以及应急预案等；而业务开发则会把更多的注意力放在业务层面的容量评估和各类策略上。同时，PE 和业务开发关注的内容是相互依赖的，他们的经验也有非常大的互补性和依赖性，所以这个过程，双方必须要紧密配合。

第三，这个过程中，你可能会发现，工具开发和稳定性开发的同学在其中参与并不多，这是因为他们的价值更多体现在平时。我们依赖的各类平台和工具，比如扩缩容、链路分析、测试数据制造、压测流量的模拟等能力，都是通过这两个团队在平时开发出来的。对于这两个团队来说，这个时候出现得越少，他们的价值才是越大的。

讲到这里，对于 SRE 体系中，每个角色要起到的作用，以及他们之间的协作机制就讲完了，各方职责不同，但是彼此互补和依赖，只有密切合作才能发挥出 SRE 体系的力量。

赛场上的哪些工作可以例行化？

谈完了“以赛带练”这种大促场景下的协作机制，那没有大促的时候，SRE 组织中的各个角色平时应该要做些什么呢？

其实，我们前面提到的“以赛带练”的事情，会有一部分转化为例行工作，同时还会增加一些周期性的工作。总结起来就是以下两项主要工作。

第一项，核心应用变更及新上线业务的稳定性评审工作。这里就包括前面讲到的容量评估和压测、预案策略是否完备等工作。PE 会跟业务开发一起评估变动的影响，比如变动的业务逻辑会不会导致性能影响，进而影响容量；对于新增加的接口或逻辑，是否要做限流、降级和熔断等服务治理策略，如果评估出来是必需的，那上线前一定要把这些策略完善好；同时在测试环境上还要做验证，上线后要关注 SLO 是否发生变化等。

第二项，周期性技术运营工作。这些就包括了我们要例行关注错误预算的消耗情况，每周或每月输出系统整体运行报表，并召集业务开发一起开评审会，对变化较大或有风险的 SLO 重点评估，进而确认是否要采取改进措施或暂停变更，以此来驱动业务开发关注和提升稳定性。

这里的技术运营工作，是 PE 职责非常大的转变，因为随着各类平台的完善，很多原来依赖运维完成的事情，业务开发完全可以依赖平台自主完成，无需运维介入。比如代码发布、配置变更，甚至是资源申请。

所以，这时我们会更强调 PE 从全局角度关注系统，除了稳定性，还可以关注资源消耗的成本数据，在稳定和效率都有保证的前提下，也能够做到成本的不断优化。这样也会使得 PE 从原有繁琐的事务中抽离出来，能够去做更有价值的事情。关于这一点，也是给运维同学一个转型和提升的建议。

总结

好了，我们做下本节课的总结。

我们借助大促这样的场景，通过“以赛带练”的思路来驱动稳定性体系的建设和提升。这个过程中，PE 更加关注全局和系统层面的稳定性，并提供各类生产环境的运行数据；而业务开发则会更关注业务代码和逻辑层面的稳定性，与 PE 互补且相互依赖；而稳定性和工具开发提供平台和工具支撑，保障每一个环节更加高效地完成。最后，我们还会将这些工作例行化，转化成日常的稳定性保障事项。

思考题

最后，给你留一个思考题。

对于“以赛带练”的思路，除了今天我们介绍到的，在你实际的工作中还有哪些场景可以尝试？

欢迎你在留言区分享，或者提出你对本节课程的任何疑问，也欢迎你把本篇文章分享给你身边的朋友。

我是赵成，我们下节课见。

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇　09 | 案例：互联网典型的SRE组织架构是怎样的?

精选留言 (2)

写留言

wholly
2020-04-08

平时以赛带练的场景有很多，除了系统集成测试方的压力测试、可靠性测试、性能专项测试外，还经常做一些局点演示及故障模拟训练，这些都是一些快速暴露问题和提取改进点的有效方式，持续提升系统稳定性。

1

leslie
2020-04-08

"以赛代练"个人觉得这种方式应当是源自体育：尤其是一些主要的比赛，我们会发现有各种各样的比赛；真正的比赛可能就是联赛，不同情况下还会有各种舍取。
电商平台中不少典型的可能会在双11之类的大型促销出现的策略，我们已经能够在平时的小型促销中去看到；合适的策略就是在生产环境的某些场景中可以利用平时业务量不大降低资源投入，非真正核心交易环节，平时降低资源投入测试一下其实风险并不大，只要...

展开